

Data-Mining für die Analyse von Nachfrage und Angebot im Revenue Management am Beispiel von Fluggesellschaften

Catherine Cleophas

Veröffentlicht in:
Multikonferenz Wirtschaftsinformatik 2012
Tagungsband der MKWI 2012
Hrsg.: Dirk Christian Mattfeld; Susanne Robra-Bissantz



Braunschweig: Institut für Wirtschaftsinformatik, 2012

Data-Mining für die Analyse von Nachfrage und Angebot im Revenue Management am Beispiel von Fluggesellschaften

Catherine Cleophas

Freie Universität Berlin, Department Wirtschaftsinformatik, 14195 Berlin,
E-Mail: catherine.cleophas@fu-berlin.de

Abstract

Nachfrageprognose und Angebotsteuerung werden im Revenue Management zunehmend auf die Betrachtung komplementäre und substituierbare Produkte ausgeweitet. Im Revenue Management bedeutet dies die Prognose und Optimierung auf der Ebene von Reisewegen, Verkaufsstandorten, Buchungsklassen und Buchungsklassengruppen. Diese detaillierte Betrachtungsweise ermöglicht eine gezielte Optimierung der Angebotssteuerung, doch sie erhöht auch die Komplexität des Modells. Dieses Paper präsentiert eine Möglichkeit zur Analyse und Klassifikation von Nachfrage und Angebot durch die Anwendung von Data-Mining. Eine entsprechende Klassifikation bildet eine Möglichkeit, das Problem kleiner Zahlen und den Aufwand komplexer Einflüsse zu reduzieren.

1 Einleitung

Revenue Management wird in [14] als ein Weg beschrieben, Informationssysteme so einzusetzen, dass die richtige Kapazität dem richtigen Kunden zur richtigen Zeit und zum richtigen Preis angeboten wird. Diese Zuordnung soll den Erlös maximieren. Für den erfolgreichen Einsatz von Revenue Management nennen [14] und [7] einige Bedingungen: dazu gehören standardisierte Produkte, eine wenig flexible Kapazitätsplanung, begrenzte Verkaufsfristen und unterschiedliche Nachfragesegmente.

Ohne eine Segmentierung der Nachfrage in Hinblick auf Zahlungsbereitschaft und Produktansprüche scheitert das Revenue Management. Eine Optimierung des Angebots auf Basis von Zielfunktion und Nebenbedingungen setzt (in der Mehrzahl der existierenden Ansätze) eine auf dieser Segmentierung basierende Prognose voraus.

Der inhaltliche Umfang des vom Revenue Management verwendeten Modells wächst ständig. Während nach [9] erste Optimierungsansätze bei Fluggesellschaften auf individuelle Flüge begrenzt waren, ist zwischenzeitlich eine netzwerkorientierte Betrachtung nach Reisewegen und Verkaufspunkten verbreitet. Die Bedingung unabhängiger Nachfrage wurde zunehmend durch Modelle beseitigt, die die Nachfrage als abhängig von der Verfügbarkeit

annehmen (siehe [16]). Aktuelle Ansätze wie zum Beispiel [7] betrachten außerdem Substitutionseffekte auf parallelen Abflügen und strategisches Kundenverhalten.

Abschnitt 2.1 zeigt, dass durch die Ausweitung des Modells die Zahl der Variablen für die Prognose und Optimierung wächst. Eine steigende Menge von Variablen hat zwei Konsequenzen: Erstens führt wie in Abschnitt 2.2 beschrieben eine feine Aggregation der Nachfrage zum sogenannten Problem der kleinen Zahlen. Zweitens bringt die Komplexität des Modells Gefahren für die Qualität der Auswertung und Beeinflussung durch Revenue Management Analysten mit sich. Diese ist allerdings wie in Abschnitt 2.3 beschrieben notwendig und wichtig, da Elemente der Unsicherheit (z.B. bedingt durch Konjunktur oder Wettbewerb) bisher nicht vollständig automatisiert berücksichtigt werden können.

Die beschriebenen Konsequenzen motivieren die hier vorgeschlagene neue Betrachtungsweise von Nachfrage und Angebot auf Basis von Techniken des Data-Minings. Abschnitt 2.4 gibt eine Einführung in den bisherigen Einsatz von Data-Mining im Operations Research im Allgemeinen und im Revenue Management im Besonderen.

Eine Beschreibung der Anwendung von Data-Mining zur Analyse und Aggregation von Nachfrage und Angebot erfolgt in Abschnitt 3. Dabei wird eine Klassifikation der Nachfrage verwendet, um die Segmentierung der Nachfrage zu realisieren. Auf Basis der Nachfrageklassifikation können Nachfrageströme auf der detailliertesten Ebene eines Modells charakterisiert und gebündelt betrachtet werden. Die sich ergebende neuartige Aggregation kann nicht nur das Problem der kleinen Zahlen reduzieren, sondern auch die Auswertung und Veränderung der Nachfrageprognose durch Analysten erleichtern. Ein ähnlicher Ansatz kann auch für die Analyse der Angebotsituation verwendet werden. Hier charakterisieren die sich aus der Optimierung ergebenden Verfügbarkeiten, Kapazitäten und die aktuelle Buchungssituation das Angebot. Eine Kombination des klassifizierten Angebots mit den zuvor erzeugten Nachfragebündeln kann nicht nur die Grundlage für die Auswertung und Beeinflussung des Angebots, sondern auch für die Erfolgsmessung bilden.

Ein Ansatz zur Validierung des hier eingeführten Konzepts wird in Abschnitt 4 vorgestellt. Eine Simulation eignet sich für die Untersuchung des Erfolgs des hier vorgestellten Konzeptes besonders, da darin die tatsächlich existierende Nachfrage transparent ist. Zusätzlich ist eine Fallstudie anhand der realen Daten einer Fluggesellschaft zur Überprüfung von Machbarkeit und Nutzen notwendig.

2 Konsequenzen des wachsenden Revenue Management Modells

Auf Basis steigender Rechnerkapazitäten und weiterentwickelter mathematischer Methoden werden die Möglichkeiten des Revenue Managements seit den siebziger Jahren ständig erweitert. Im Prozess des Operations Research wie in [10] beschrieben bildet ein Entscheidungsmodell die Grundlage für die Suche nach einer optimalen Lösung. Im Revenue Management sind in diesem Entscheidungsmodell Annahmen über das Verhalten von Kunden enthalten. Durch eine Annäherung des Modells an die Realität soll, als Grundlage für die Angebotssteuerung, eine ideale Segmentierung vorgenommen werden.

Dieser Abschnitt dokumentiert einige Schritte in dieser Entwicklung und zeigt, dass die Erweiterung des Modells auch eine Vermehrung der darin enthaltenen Variablen und damit der hierarchischen Komplexität zur Folge hat. Die Konsequenzen aus einer Vielzahl von

Variablen in Form des Problems kleiner Zahlen und eines steigenden Aufwandes bei der Bewertung und Beeinflussung werden erläutert. Das Data-Mining wird als Methode zum Umgang mit großen und komplexen Datenmengen eingeführt. In diesem Zusammenhang werden schließlich existierenden Anwendungen von Data-Mining in Operations Research und Revenue Management zusammengefasst.

2.1 Entwicklung des Revenue Management Modells

Mit wenigen Ausnahmen basiert Revenue Management nach [14] auf einer Kombination von Prognose und Optimierung. Eine Nachfrageprognose parametrisiert Nebenbedingungen, welche die Grundlage der Optimierungsfunktion bilden. Prognose und Optimierung verwenden ein gemeinsames Modell, welches den Grad der Nachfrageaggregation ebenso wie die Ebene, auf der die Verfügbarkeit berechnet wird, bestimmt.

Herausforderungen wie in [3] beschrieben können häufig als Erweiterungen des mathematischen Modells zur Annäherung an die Realität formuliert werden. Indem bisherige Beschränkungen überwunden und neue Methoden entwickelt werden, wächst das Modell. Die sich daraus ergebenden Effekte werden unter anderem auch in [1] demonstriert.

Im Revenue Management von Fluggesellschaften konzentrieren sich nach [9] die ersten Ansätze auf die Prognose und Optimierung von einzelnen Flügen unter der Annahme unabhängiger Nachfrage. In diesem Modell gibt es keine Unterschiede in der Bewertung von Umsteigern (die zwei komplementäre Produkte kaufen) und lokalen Passagieren (die nur eines kaufen). Weiterhin gibt es keine Substitute, so dass die Verfügbarkeit einer Buchungsklasse keinen Einfluss auf die Nachfrage nach anderen Buchungsklassen hat. Das flugbasierte Modell führt bei einem Netzwerk aus acht Flügen und vier angebotenen Buchungsklassen, unabhängig von der Netzstruktur, zu $8 * 4 = 32$ Variablen. Für jede dieser Variablen wird die Nachfrage prognostiziert und die erlösoptimale Verfügbarkeit berechnet.

Eine erste Erweiterung erfolgte mit der Betrachtung der durch die Kombination einzelner Flüge gebildeten Reisewege. Hier wird die Nachfrage für komplementäre Güter zusammen prognostiziert, um durch die gemeinsame Optimierung den Erlös zu steigern. Beispiele für netzwerkorientiertes Revenue Management finden sich in [11].

Geht man, wie im oben genannten Beispiel, weiterhin von acht Flügen und vier angebotenen Buchungsklassen aus, hängt die Zahl der sich ergebenden Variablen eines netzwerkorientierten Modells von der Struktur des Netzwerks ab. Wir betrachten ein kleines Netz aus drei Flughäfen: jeweils zwei Flüge verbinden Spoke 1 mit dem Hub und den Hub mit Spoke 2 in beide Richtungen. Zwei der drei Flughäfen sind also nicht durch einen direkten Flug verbunden. In diesem Netzwerk ist die Anzahl der Reisewege maximal 16, wenn die Anschlusszeit für alle sinnvollen Verbindungen passend ist. Dabei wird angenommen, dass kein Reiseweg denselben Start- und Endpunkt hat und keine unnötigen Umwege in Kauf genommen werden. Stellen die Knoten des Netzwerks unterschiedliche Verkaufspunkte dar, und kann jeder Reiseweg von jedem Verkaufspunkt aus gebucht werden, ergeben sich $16 * 3 = 48$ mögliche Kombinationen von Reisewegen und Verkaufspunkten. Bei vier Buchungsklassen ergeben sich $4 * 16 * 3 = 192$ Variablen.

Eine jüngere Ergänzung des Modells ist die Betrachtung von Substitutionseffekten zwischen zeitgleich angebotenen Kombinationen von Produkten und Preisen. Für Fluggesellschaften

ist die Nachfrage abhängig von der Verfügbarkeit paralleler Flüge und anderer Buchungsklassen, wie in [16] beschrieben.

Besteht die Möglichkeit, dass Kunden mehrere Klassen akzeptieren, muss die Nachfrage auch für denkbare Kombinationen von Substituten prognostiziert werden. Bei vier Buchungsklassen ergeben sich maximal vier Kundengruppen, die nur eine Klasse akzeptieren, sechs Gruppen à zwei Klassen, vier Gruppen à drei Klassen und eine Gruppe, die alle Klassen akzeptiert. Für jeden Reiseweg und jeden Verkaufspunkt müssen demnach 15 Kundengruppen prognostiziert und in der Optimierung berücksichtigt werden. Dies erhöht die Anzahl der Variablen auf $15 * 16 * 3 = 720$ Variablen.

Bisher wurden Nachfrage und Angebot nicht über den Verkaufszeitraum zeitlich differenziert prognostiziert und optimiert. Eine Aufteilung des Verkaufszeitraumes in mehrere Abschnitte ist allerdings im Revenue Management weitverbreitet. Der Zeitpunkt der Nachfrage hat eine besondere Bedeutung, da oft frühzeitige Anfragen nach günstigen Angeboten erfüllt werden, gleichzeitig jedoch Kapazitäten für späte, hochwertige Anfragen reserviert bleiben sollen [8].

Wird der Verkaufszeitraum unterteilt, erhöht sich die Anzahl der Variablen für die Prognose und darauf basierend die Angebotssteuerung weiter. Wenn der Verkaufszeitraum von einem Jahr im Beispiel des Airline Revenue Managements in 20 Zeitabschnitte aufgeteilt wird, erhöht sich die maximale Zahl der Variablen im netzbasierten Modell auf $192 * 20 = 3.840$. Im netzbasierten Modell mit abhängiger Nachfrage können bis zu 14.400 Variablen entstehen.

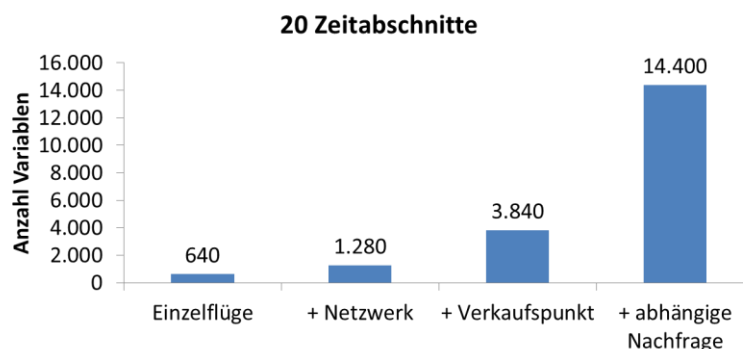


Bild 1: Anzahl der Variablen beim Ausbau des Modells

Bild 1 fasst die beschriebene Entwicklung zusammen. Das Modell wächst noch weiter, wenn die Möglichkeit der Substitution zwischen Abflugzeiten oder strategische Kunden einbezogen wird. Strategische Kunden (siehe [6]) fragen im Laufe des Verkaufszeitraums mehrfach an und fällen die Entscheidung, das Produkt zum angebotenen Preis zu kaufen oder weiter abzuwarten, basierend auf ihren Erfahrungen über den Preisverlauf. Die Kunden können also nicht nur zwischen Buchungsklassen substituieren, sondern auch die Entscheidung über den Zeitpunkt ihrer wiederholten Abfrage zur Preisminimierung nutzen. Eine weitere Betrachtung der Konsequenzen des Modellwachstums für das Revenue Management unter dem Aspekt der Komplexität findet sich in [1].

2.2 Das Problem der kleinen Zahlen

Der mit dem Wachstum des Modells verbundene Aufwand kann durch den methodischen und technischen Fortschritt teilweise kompensiert werden. Doch durch die Vielzahl der Variablen in der Nachfrageprognose entsteht auch das sogenannte Problem der kleinen Zahlen. Dieses ergibt sich, wenn ein Modell mit einer Vielzahl von Parametern durch eine kleinere Zahl von Beobachtungen definiert werden soll. Im Revenue Management kann die Zahl der entstehenden Variablen die zu einem Zeitpunkt angebotene Kapazität und damit die Anzahl der möglichen beobachteten Käufe weit übersteigen. Dieses Problem wird unter anderem in [2] kurz erläutert.

Das Problem kleiner Zahlen tritt bereits für das im vergangenen Abschnitt beschriebene Beispiel schnell auf. Nimmt man 160 Sitzplätze pro Flug an, ergeben sich maximal 1.280 verkaufte Tickets ohne Überbuchung. Nach dieser Rechnung kommen bei 3.840 Variablen auf jede Beobachtung drei Variablen.

Durch eine Akkumulation der Beobachtungen über mehrere Abflüge hinweg kann diese Quote verbessert werden. Allerdings werden sich, insbesondere bei einer unregelmäßigen Verteilung von Buchungen über Zeitabschnitte, Erwartungswerte von weniger als einer Buchung ergeben. Dies tritt auf, wenn zum Beispiel eine Klasse auf einem Reiseweg und einem Verkaufspunkt zu einem Zeitpunkt nur für einen von zehn beobachteten Abflügen verkauft wurde. Diese Bruchteile können durch Rundungseffekte zu stark schwankenden Ergebnissen führen: Bereits eine zusätzliche Buchung kann zu einer Differenz zwischen tatsächlicher und prognostizierter Nachfrage von über 100 % führen.

Wie auch in [2] erläutert ist es nicht sinnvoll, Reisewege mit kleinen Verkaufszahlen zu ignorieren. Während die Nachfrage im Einzelnen gering sein mag, kann sie in der Summe einen wesentlichen Anteil am gesamten Erlös haben. Es gibt also einen Bedarf nach Ansätzen zur Aggregation, die es gestatten, die Zahl der Parameter zu verringern ohne den für die sinnvolle Segmentierung notwendigen Informationsgehalt aufzugeben.

2.3 Analyse und Beeinflussung komplexer Modelle

Trotz der Weiterentwicklung des Revenue Management Modells bleiben Elemente der Unsicherheit bestehen, die mittelfristig nicht durch automatisierte Algorithmen berücksichtigt werden können. Beispiele sind Veränderungen der Nachfrage durch starke konjunkturelle Schwankungen und Wettbewerberstrategien. In [11] wird gezeigt, dass die Pflege der Prognose durch Analysten zur Erlösmaximierung beiträgt. Dies ist unter anderem notwendig, wenn sich das Nachfrageverhalten kurzfristig ändert. Nach [5] darf ein vollständig automatisiertes wettbewerbsorientiertes Revenue Management aus spieltheoretischer Sicht wenig Erfolg erwarten. Die Autoren überlassen es der Intuition der Analysten, die Erlössteuerung auf die Existenz und das Verhalten des Wettbewerbers anzupassen.

Analysten können nicht ohne einen erheblichen Mehraufwand Auswertungen und Einflüsse auf der Ebene realisieren, auf der automatisierte Systeme arbeiten. Somit wird eine Aggregation und Darstellung benötigt, die die Auswertung auf einer weniger detaillierten Ebene gestattet und gleichzeitig den Informationsverlust minimiert. Auf einer derartigen Sicht auf das System kann eine unmittelbare Beeinflussung der Ergebnisse aufsetzen.

2.4 Data-Mining und Revenue Management

Data-Mining ist nach [12] als „automatisierter oder semi-automatisierter Prozess zur Extraktion bisher unbekannten und potentiell nützlichen Wissens aus großen Datenbanken“ definiert. Dies geht über die systematische Anwendung von Datenhaltung und -verarbeitung hinaus und muss davon abgegrenzt werden. Wie in [10] beschrieben kann die Effektivität des Operations Research durch Data-Mining unterstützt werden. Die Anwendung von Data-Mining auf das Revenue Management bietet dafür eine Gelegenheit. Dabei zielt die hier beschriebene Methode nicht auf eine Ablösung der herkömmlichen Methodik durch Data-Mining („Effektivität durch Ersatz“ nach [10]) sondern auf eine Verbesserung durch Kombination („Effektivität durch Verfeinerung“ [10]) ab.

Die Anwendung des Data-Mining zur Unterstützung des Revenue Managements ist bisher kaum dokumentiert. Die Möglichkeiten zur Verwendung von Techniken des Data-Mining im Bereich Customer Relationship Management, wie sie [14] zusammenfasst, sind allerdings auf die bisher beschriebenen Problemstellungen des Revenue Managements übertragbar. In beiden Fällen ist das Ziel eine Segmentierung der Kunden auf Basis historischer Daten. Wie zum Beispiel in [13] vorgeschlagen können Muster im Kundenverhalten mithilfe von Data-Mining identifiziert, standardisiert und für die Klassifikation von neuen Beobachtungen verwendet werden. Der wesentliche Unterschied besteht darin, dass das Customer Relationship Management Daten in Bezug auf das Verhalten individueller Kunden sammelt, während im Revenue Management üblicherweise Daten auf der feinsten im verwendeten Modell gegebenen Ebene gesammelt werden.

Revenue Management und Customer Relationship Management rücken noch näher aneinander, wenn der langfristige Kundenwert in Betracht gezogen wird. Eine derartige Betrachtung wird u.a. in [17] vorgeschlagen. Der weitere Text ist auf die Identifikation von Kundensegmenten für das netzorientierte Revenue Management Modell mit abhängiger Nachfrage beschränkt. Weitere Aspekte, wie der langfristige Kundenwert oder auch strategisches Kundenverhalten, bieten eine zusätzliche Motivation für Weiterentwicklungen.

3 Data-Mining zur Analyse von Nachfrage und Angebot

Data-Mining kann in Revenue Management an mehreren Stellen wie in Bild 2 illustriert angewendet werden. Erstens kann es die Segmentierung der Nachfrage unterstützen. Zweitens kann Data-Mining helfen, das Problem kleiner Zahlen zu begrenzen, indem ähnliche Instanzen identifiziert und aggregiert prognostiziert werden. Drittens kann es die Analyse und Beeinflussung der Nachfrageprognose und Angebotssteuerung unterstützen.

Auf diese Weise werden drei der in [10] beschriebene Anwendungsarten des Data-Mining genutzt: Die *Beschreibung* von Ähnlichkeiten zwischen Instanzen etc., die *Mustererkennung* zur Erklärung dieser Ähnlichkeiten und die *prädiktive Folgerung* aus ähnlichen Attributen der Instanzen auf eine ähnliche Entwicklung der Nachfrage. Bild 3 illustriert Gelegenheiten zur Einbindung von Data-Mining in den Revenue Management Prozess.

In den folgenden Abschnitten wird die Verwendung von Data-Mining zur Segmentierung der Nachfrage, Klassifikation von Produktkombinationen und Identifikation von Erlöspotentialen beschrieben. Die im weiteren Text verwendeten Definitionen und Methoden des Data-Mining

sind, sofern nicht ausdrücklich anders identifiziert, aus [18] übernommen. Ihre Anwendung auf das Revenue Management ist ein in diesem Beitrag neu vorgestellter Ansatz.

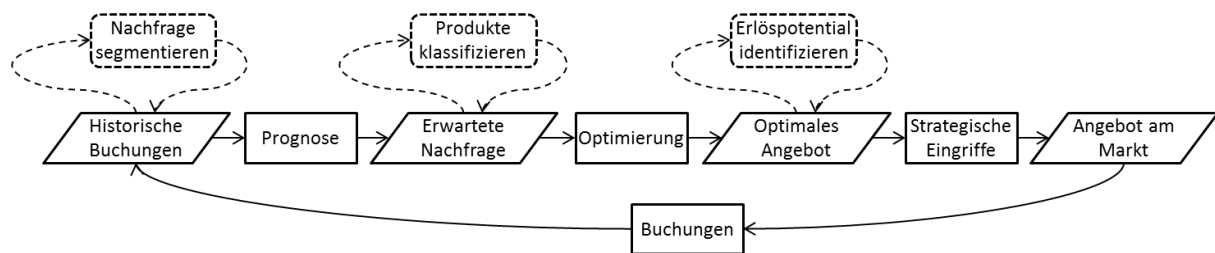


Bild 2: Data-Mining im Revenue Management Prozess

3.1 Segmentierung der Nachfrage

Im traditionellen Revenue Management basiert die Nachfrageprognose auf historischen Buchungen. Dies wirft das hier nicht weiter behandelte Problem des Unconstrainings auf – Buchungen, die durch die Angebotskonstellation begrenzt sind, müssen in die ursprüngliche Nachfrage transformiert werden. Eine Erläuterung dieser Problematik sowie ein Vergleich möglicher Lösungsansätze findet sich in [15]. Historische Buchungen sind durch folgende Merkmale charakterisiert: gekauftes Produkt, Zeitpunkt des Kaufs, Verkaufsweg, Angebotskonstellation zum Zeitpunkt des Kaufs, Preis. Betrachtet man den Einsatz bei Fluggesellschaften, können diese Merkmale wie folgt weiter präzisiert werden.

Das *gekaufte Produkt* ist der Transport über einen Reiseweg (Zeitpunkt von Abflug und Ankunft als Zeit und Datum, Ausgangsort, Zielort und ggf. Umsteigeorte, Reisezeit) zu den Restriktionen des gebuchten Tarifs. Der *Zeitpunkt des Kaufs* kann in Zeitabschnitten vor Abflug oder als vom Abflug unabhängige Kombination aus Uhrzeit, Wochentag und Datum beschrieben werden. Auch der *Verkaufsweg* kann unterschiedlich beschrieben werden. Er kann als Start oder Ziel der Reise oder als geographische Einheit angegeben werden. Auch eine Unterscheidung zwischen dem Verkauf über Reiseagenten, Firmenbüros und Internetanbieter ist möglich. Die *Angebotskonstellation* spielt dann eine Rolle, wenn der weitere Prozess abhängige Nachfrage in Betracht zieht. Sie kann als die Verfügbarkeit günstigerer Tarife auf demselben oder vergleichbaren Reise- und Verkaufswegen beschrieben werden, aber auch zeitgleich bestehende Angebote von Wettbewerbern enthalten. Weiterhin kann zwischen dem *Preis* abzüglich von Steuern und Gebühren, inklusive Steuern und Gebühren oder auch inklusiver zusätzlicher Kosten zum Beispiel für Gepäckstücke unterschieden werden.

Das Ergebnis einer derartigen Beschreibung von historischen Buchungen ist ein Datensatz pro Buchung. Das Data-Mining bezeichnet diesen Datensatz als *Instanz*, die beschriebenen Merkmale als *Attribute*. Diese Begriffe werden im weiteren Text in diesem Sinne verwendet. Die Instanzen können *informiert* oder *uninformiert* segmentiert werden – die verschiedenen bringen verschiedene Anforderungen und Möglichkeiten mit sich.

Das *Clustering* stellt eine uninformierte Methode der Segmentierung dar. Der Vorteil des uniformierten Lernens von Kundensegmenten ist die ergebnisoffene Vorgehensweise: Die entstehenden Cluster sind von bestehenden Hypothesen unabhängig und können neue, unerwartete Einblicke in Zusammenhänge zwischen den Attributen bieten. Gleichzeitig sind die entstehenden Cluster nicht notwendigerweise im weiteren Prozess nutzbar, da sie sich

nicht unbedingt an Attributen orientieren, die traditionell im Revenue Management betrachtet werden.

Die *Klassifikation durch Entscheidungsbäume* ist ein Beispiel für die informierte Segmentierung. Dafür müssen Klassen vorab auf Basis von Attributausprägungen definiert werden. Für Fluggesellschaften können beispielsweise Beschreibungen von Kundensegmenten anhand der Attribute gezahlter Preis, gebuchter Tarif und Anfragezeitpunkt für die Definition der Klassen genutzt werden. Diese Methode hat den Vorteil, dass die Ergebnisse mit dem verwendeten Revenue Management Modell unmittelbar verbunden werden können. Ein zusätzliches, der Klassifikation vorausgehendes Clustering kann bei der Suche nach besonders vielversprechenden Attributen nützlich sein.

Die Ergebnisse dieses Schritts können unabhängig zur Analyse der Nachfrage, aber auch für den Vergleich von beobachteter und prognostizierter Nachfrage verwendet werden. Entstehende Klassen und Regeln können eine Grundlage für den Vergleich von Märkten sowie für die Analyse von Marktentwicklungen über die Zeit bilden. Die weiteren hier beschriebenen Schritte gehen von einer Klassifikation nach gezahltem Preis, gleichzeitig verfügbaren Buchungsklassen und vom Kunden gewählten Tarifeigenschaften aus.

3.2 Bündelung von Produktkombinationen

Abhängig von den betrachteten Attributen und dem Kundenmodell der Prognose kann die in Abschnitt 3.1 beschriebene Segmentierung der Nachfrage sowohl auf die historischen Buchungen als auch auf die prognostizierte Nachfrage angewendet werden. Jede Produktkombination lässt sich durch den Anteil der dafür in der Vergangenheit beobachteten und prognostizierten Nachfragesegmente beschreiben. Im Kontext von Fluggesellschaften kann eine Produktkombination durch den Reiseweg, den Verkaufsweg und den Zeitabschnitt des Verkaufszeitraums beschrieben werden. Außerdem können Besonderheiten des angebotenen Tarifs, wie zum Beispiel die Unterscheidung zwischen Business- und Economy-Klassen, berücksichtigt werden.

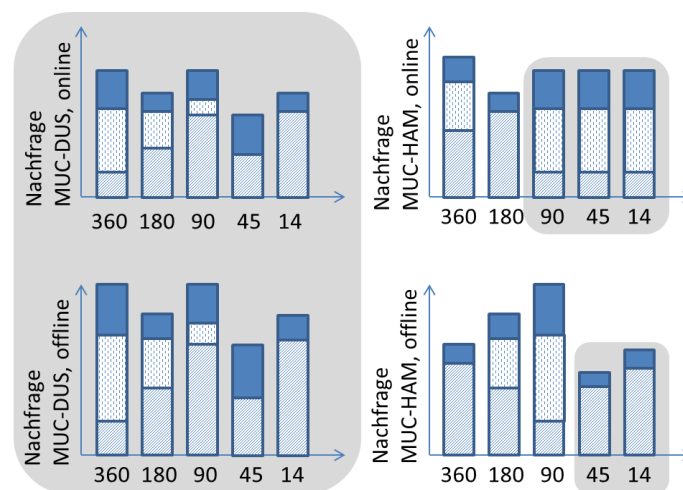


Bild 3: Bündelung von Produktkombinationen anhand von Nachfragesegmenten

Während im vorausgehenden Abschnitt einzelne Buchungen als zu klassifizierende Instanzen betrachtet wurden, dienen nun Produktkombinationen als Instanzen. Bild 3 zeigt ein beispielhaftes Ergebnis auf jeweils zwei Reise- und Verkaufswegen. Die Kundenmenge

pro Zeitabschnitt des Verkaufshorizontes ist als Säule abgetragen. Dabei ist jede Säule in die Anteile der für diesen Abschnitt erwarteten Kundensegmente aufgeteilt.

Im Anschluss an eine derartige Charakterisierung können ähnliche Instanzen zusammengefasst werden. In Bild 3 wird die Ähnlichkeit der Verkaufswege *online* und *offline* auf dem Reiseweg *MUC-DUS*, sowie die Ähnlichkeit einiger aufeinander folgender Zeitabschnitte im Verkaufshorizont des Reisewegs *MUC-HAM* betont.

Die Identifikation ähnlicher Produktkombinationen auf Basis der Beschreibung durch Kundensegmente kann durch Clustering unterstützt werden. Zusätzliche Informationen, wie in der Vergangenheit beobachtete Trends in der Menge der Nachfrage, für eine Kombination können als Attribute einbezogen werden.

Die Plausibilität und Akzeptanz des Ergebnisses wird verbessert, wenn Experten die Randbedingungen der Identifikation ähnlicher Produktkombinationen festlegen können. Eine mögliche Randbedingung ist die Beschränkung der Bündelung auf aufeinanderfolgende Zeitabschnitte oder Reisewege zwischen übereinstimmenden Ländern.

Die resultierenden Produktbündel können zur Reduktion des Problems kleiner Zahlen und der Auswertung und Beeinflussung der Prognose durch Analysten genutzt werden. Das Problem kleiner Zahlen wird beschränkt, indem für ähnliche Bündel eine gemeinsame Prognose erstellt wird. Da die Beobachtungen für das gesamte Bündel genutzt werden, ergeben sich größere Zahlen. Die so erstellte Prognose kann entweder auf die einzelnen Produktkombinationen umgelegt oder für eine gemeinsame Optimierung verwendet werden.

Bei der Auswertung sind zwei Verwendungszwecke zu unterscheiden. Einerseits kann die prognostizierte Nachfrage mit der Einschätzung des Analysten verglichen werden. Sind die im Bündel enthaltenen Produktkombinationen zum Beispiel von einer Messe betroffen, kann die erwartete Auswirkung vom Analysten in die Prognose eingepflegt werden. Abhängig von der Einschätzung des Analysten ist es auch möglich, Produktkombinationen aus einem Bündel mit den Eigenschaften eines anderen Bündels zu versehen, wenn eine Berichtigung der automatisierten Prognose in diesem Maß notwendig erscheint. Andererseits können nach Ende des Verkaufszeitraums Vergleiche zwischen der prognostizierten und der tatsächlich beobachteten Nachfrage aufgestellt werden, um die Qualität der Prognose zu beurteilen. Eine gute Prognose sollte die tatsächlich beobachtete Nachfrage als Teilmenge enthalten, aber nicht auf sie beschränkt sein.

3.3 Identifikation von Erlöspotentialen

Die dritte Gelegenheit, Data-Mining zur Unterstützung des Revenue Managements einzusetzen, ergibt sich nach Abschluss der Optimierung. Durch ihre Abhängigkeit von Prognose und Kapazität sind die Auswirkungen der Optimierung anhand einer herkömmlichen Analyse der Verfügbarkeitskonstellation nicht unbedingt offensichtlich. Zudem kann ein differenziertes Revenue Management Modell zu Mustern in den Erlöspotentialen führen, die ohne eine systematische Analyse auf der kleinsten Ebene nicht zu erkennen sind. Eine derartige Analyse ist aufgrund der Vielzahl der Variablen, wie in Abschnitt 2.1 dargestellt, für Analysten anspruchsvoll und aufwändig.

Wäre die automatisierte Optimierung der letzte Schritt des Revenue Management Prozesses, könnte auf diese Analyse gegebenenfalls verzichtet werden. Doch strategische Nebenziele wie zum Beispiel die flexible Reaktion auf Wettbewerberangebote erfordern ein

tiefgehendes Verständnis der Situation und eine Möglichkeit zu gezielten Eingriffen durch Analysten (siehe [5]).

Die Identifikation von Erlöspotentialen kann auf der Ebene individueller Produktkombinationen oder der in Abschnitt 3.2 beschriebenen Bündel stattfinden. Abhängig davon werden entweder einzelne Produktkombinationen oder Bündel in diesem Schritt als Instanzen behandelt.

Anhand der nach der Optimierung verfügbaren Informationen können neben der prognostizierten Nachfrage weitere Merkmale identifiziert und zur Klassifikation der Instanzen verwendet werden. Insbesondere der *maximale Anteil der Gesamtkapazität*, der erwartete *durchschnittliche Erlös pro Einheit* sowie die erwartete *Produktivität als Prozentsatz der Kapazität* sind denkbare Attribute. Auch Informationen über die Wettbewerbssituation und die Existenz besonderer Ereignisse können miteinbezogen werden. Diese Attribute können gemeinsam mit Attributen, die die Zusammensetzung der Nachfrage beschreiben, für die Klassifikation der Instanzen verwendet werden.

Die Beschreibung der verwendeten Klassen hängt von der Zielsetzung dieses Schritts ab. Sollen primär die für den Erlös vielversprechendsten Instanzen identifiziert werden, bietet sich eine Einteilung nach erwarteter Produktivität und durchschnittlichem Erlös an. Sind Nebenziele wie eine möglichst gute Ausnutzung der gegebenen Kapazität oder eine wettbewerbsorientierte Steuerung die Motivation, können die verwendeten Klassen die verfügbare Kapazität oder Abstufungen der Wettbewerbsintensität miteinbeziehen.

Das Ergebnis dieses Schrittes kann auf mindestens zweierlei Weise verwendet werden. Einerseits beschreiben die entstandenen Klassen die enthaltenen Instanzen und geben Analysten so einen Überblick über die Anteile und die Verteilung von hochwertigen Produkten am gesamten Angebot. Auf Basis dieses Überblicks können zum Beispiel Einflüsse mit dem Ziel, die Produktivität zu steigern, gesetzt werden.

Andererseits bietet die Klassifikation auch eine neue Möglichkeit zur Erfolgsmessung im Revenue Management. Durch die sämtliche der beschriebenen Schritte hinweg kann eine Topologie von Erlöspotentialen aufgebaut werden, die auf Attributen der Nachfrage und der Optimierung basiert. Nach Ende des Verkaufszeitraumes wäre zu erwarten, dass die durch Prognose und Optimierung beschriebene Situation in Erlöse umgesetzt wurde. Zum Beispiel sollte ein Reiseweg mit hochwertiger prognostizierter Nachfrage und hohem angestrebten Durchschnittserlös bei einer Berechnung des erzielten Erlöses weiterhin als hochwertig zu erkennen sein. Ist dem nicht so, muss der Grund entweder in der Prognose, der automatisierten Optimierung oder der strategischen Steuerung gesucht werden.

4 Geplante Evaluation durch Simulation und Fallstudie

Der beschriebene Ansatz zur Nutzung von Data-Mining für die Analyse von Nachfrage und Angebot im Revenue Management kann wie in [4] beschrieben durch den Einsatz von Simulationen evaluiert werden. Simulationen gestatten es, Erkenntnisse des Data-Minings mit dem vorhandenen Wissen über die tatsächliche Nachfrage zu vergleichen: Eine Simulation für das Revenue Management enthält individuelle, künstlich generierte Kunden, deren Verhalten vorab ausgewertet werden kann. Mithilfe des entstehenden Wissens kann der Erfolg des Data-Minings in jedem Schritt evaluiert werden.

Zusätzlich müssen aus den hier beschriebenen Konzepten hervorgehende konkrete Vorhaben und Methoden mit realen Daten konfrontiert werden. Die Menge und Qualität der in der Praxis vorhandenen Daten und die zur Auswertung zur Verfügung stehende Zeit muss bei der Entwicklung der Methodik in Betracht gezogen werden.

Eine Überprüfung des hier vorgestellten Konzeptes ist in Zusammenarbeit mit einer großen deutschen Fluggesellschaft geplant. Im diesem Rahmen können die drei gezeigten Schritte sequentiell durchgeführt und gemeinsam mit Experten bewertet werden. Dabei erfolgt die Segmentierung der beobachteten Nachfrage auf Basis der historischen Buchungsdaten auf Kombinationen aus Reisewegen, Verkaufspunkten, Verkaufszeitabschnitten, Tarifkombinationen und Preisen über mehrere Zeitabschnitte. Anschließend kann die Klassifikation von Produktkombinationen umgesetzt werden. Im Kontext der Fluggesellschaft bedeutet dies eine Einteilung von Kombinationen aus Reisewegen, Verkaufspunkten und Tarifbedingungen nach dem darin beobachteten Nachfrageverhalten. Schließlich stellt die Identifikation von Erlöspotentialen im Kontext der Fluggesellschaft die aufwändigste, gleichzeitig aber auch die mit dem größten praktischen Interesse verbundene Aufgabe dar.

5 Zusammenfassung und Ausblick

In diesem Text wurde die Notwendigkeit einer verbesserten Auswertung von Nachfrage und Angebot im Revenue Management aufgrund des Problems kleiner Zahlen und der Herausforderungen bei der Analyse und Beeinflussung komplexer Modelle erläutert. Als Lösungsansatz wurden drei Gelegenheiten für den Einsatz von Data-Mining im Revenue Management Prozess identifiziert und beschrieben. Die Bedingungen und Potentiale jedes Schrittes wurden insbesondere für den Einsatz bei Fluggesellschaften erläutert.

Jede der hier gezeigten Einsatzmöglichkeiten bietet bereits bei isolierter Betrachtung Vorteile. Beim sequentiellen Einsatz kann auf den Ergebnissen der vorausgehenden Schritte aufgebaut werden, was die Effektivität des Ansatzes steigern sollte. Insbesondere die Erfolgsmessung von Prognose und Optimierung stellt eine vielversprechende Einsatzmöglichkeit dar.

Das vorgestellte Konzept bedarf für jeden Schritt, wie in Abschnitt 4 skizziert, eine Validierung durch Simulationsexperimente und Fallstudien. Dabei ist eine wesentliche Aufgabe die Überprüfung der Annahme, dass die vom Revenue Management betrachteten Instanzen sich in aussagekräftige Segmente gliedern lassen. Weiterhin muss die Durchführbarkeit des Vorgehens mit Blick auf den Umfang und die Qualität realer Daten geprüft werden. Insbesondere bei der Verwendung externer Daten, zum Beispiel zur Beschreibung der wirtschaftlichen Lage, muss die Verfügbarkeit auf der benötigten Detailebene geprüft werden.

6 Literatur

- [1] Bartke, P; Cleophas, C; Zimmermann, B (im Erscheinen): Complexity in Airline Revenue Management. Journal of Revenue & Pricing Management.
- [2] Boyd, E; Bilegan, I (2003): Revenue management and e-commerce. Management Science 49(10):1363-1386.

- [3] Chiang, W; Chen, JH; Xu, X (2007): An overview of research on revenue management: current issues and future research. *International Journal of Revenue Management* 1(1):97-128.
- [4] Cleophas, C; Frank, M; Kliwer, N (2009): Simulation-based key performance indicators for evaluating the quality of airline demand forecasting. *Journal of Revenue and Pricing Management* 8(4):330-342.
- [5] Isler, K; Imhof, H (2008): A game theoretic model for airline revenue management and competitive pricing. *Journal of Revenue & Pricing Management* 7(4):384-396.
- [6] Jerath, K; Netessine, S; Veeraraghavan, S (2010): Revenue management with strategic customers: Last-minute selling and opaque selling. *Management Science* 56(3):430-448.
- [7] Klein, R; Steinhardt, C (2008): *Revenue Management – Grundlagen und mathematische Methoden*. Springer-Verlag, Berlin Heidelberg.
- [8] Littlewood, K (2005): Special Issue Papers: Forecasting and control of passenger bookings. *Journal of Revenue & Pricing Management* 4(2):111-123.
- [9] McGill, J; van Ryzin, G (1999): Revenue management: Research overview and prospects. *Transportation Science* 33(2):233-256.
- [10] Meisel, S; Mattfeld, D (2009): Synergies of operations research and Data-Mining. *European Journal of Operational Research*.
- [11] Mukhopadhyay, S; Samaddar, S; Colville, G (2007): Improving Revenue Management Decision Making for Airlines by Evaluating Analyst-Adjusted Passenger Demand Forecasts. *Decision Sciences* 38(2):309-327.
- [12] Olafsson, S; Li, X; Wu, S (2006): Operations research and Data-Mining. *European Journal of Operational Research* 187(3):1429-1448.
- [13] Padmanabhan, B; Tuzhilin, A (2003): On the use of optimization for Data-Mining: Theoretical interactions and eCRM opportunities. *Management Science* 49(10): 1327-1343.
- [14] Talluri, K; van Ryzin, G (2004): *Theory and practice of revenue management*. Kluwer Academic Publishers, Boston.
- [15] Weatherford, L; Pölt, S (2002): Better unconstraining of airline demand data in revenue management systems for improved forecast accuracy and greater revenues. *Journal of Revenue & Pricing Management* 1(3):234-254.
- [16] Weatherford, LR; Ratliff, RM (2010): Review of revenue management methods with dependent demands. *Journal of Revenue & Pricing Management* 9(4):326-340.
- [17] Wirtz, J; Kimes, S; Theng, JP; Patterson, P (2003): Revenue management: resolving potential customer conflicts. *Journal of Revenue & Pricing Management* 2(3):216-226.
- [18] Witten, I; Frank, E; Hall, MA (2011): *Data-Mining: Practical Machine Learning Tools and Techniques*. 3. Auflage. Morgan Kaufmann, Amsterdam.